Journal of Biomolecular NMR (2021) 75:221–232
https://doi.org/10.1007/s10858-021-00370-0

ARTICLE

Optimized precursor to simplify assignment transfer between backbone resonances and stereospecifically labelled valine and leucine methyl groups: application to human Hsp90 N-terminal domain

Faustine Henot[1] · Rime Kerfah[2] · Ricarda Törner[1] · Pavel Macek[1,2] · Elodie Crublet[2] · Pierre Gans[1] · Matthias Frech[3] · Olivier Hamelin[4] · Jerome Boisbouvier[1]

Received: 26 February 2021 / Accepted: 20 May 2021 / Published online: 27 May 2021

Abstract

Methyl moieties are highly valuable probes for quantitative NMR studies of large proteins. Hence, their assignment is of the utmost interest to obtain information on both interactions and dynamics of proteins in solution. Here, we present the synthesis of a new precursor that allows connection of leucine and valine pro-*S* methyl moieties to backbone atoms by linear ^{13}C-chains. This optimized ^{2}H/^{13}C-labelled acetolactate precursor can be combined with existing ^{13}C/^{2}H-alanine and isoleucine precursors in order to directly transfer backbone assignment to the corresponding methyl groups. Using this simple approach leucine and valine pro-*S* methyl groups can be assigned using a single sample without requiring correction of ^{1}H/^{2}H isotopic shifts on ^{13}C resonances. The approach was demonstrated on the N-terminal domain of human HSP90, for which complete assignment of Ala-β, Ile-δ_1, Leu-δ_2, Met-ε, Thr-γ and Val-γ_2 methyl groups was obtained.

Keywords NMR · Assignment · Methyl groups · Acetolactate · HSP90

Introduction

Solution state NMR is the method of choice to characterize proteins at atomic level and to probe their dynamics over a wide range of biologically relevant timescales. However, for a long-time, study of high molecular weight proteins by NMR remained a challenge, notably due to the extensive line broadening of NMR signals in large proteins. Methyl groups have been widely studied and are extremely useful to overcome this issue that has hampered, in the past, quantitative NMR studies on large proteins. Indeed, due to the proton multiplicity and their favorable relaxation properties, methyl groups allow the detection of NMR signals even for large proteins (Tugarinov et al. 2003). Nowadays, methyl groups are important probes to investigate molecular dynamics (Sprangers and Kay 2007) and to provide functional insight (Rosenzweig et al. 2013; Mas et al. 2018) on assemblies weighing up to 1 MDa. Specific labelling of methyl groups on perdeuterated large proteins allows the measurement of long-range distance restraints, up to 12 Å. (Sounier et al. 2007; Ayala et al. 2020) and enables, in combination with other structural biology techniques such as SANS/SAXS (Lapinaite et al. 2013) or Cryo-EM (Gauto et al. 2019), to solve the structure of complexes of several hundreds of kDa.

For the past 20 years, a plethora of protocols overexpressing proteins in M9/$^{2}H_2O$ based *E. coli* growth medium and leading to specific protonation of methyl groups in perdeuterated proteins without scrambling of protons to other sites have been elaborated. On one hand, the direct incorporation of the methyl labelled amino acid in M9/$^{2}H_2O$ is employed for the specific labelling of $^{13}C^{1}H_3$-alanine (Isaacson et al. 2007; Ayala et al. 2009), $^{13}C^{1}H_3$-methionine (Gelis et al. 2007; Stoffregen et al. 2012) and $^{13}C^{1}H_3$-threonine (Velyvis

✉ Jerome Boisbouvier
jerome.boisbouvier@ibs.fr

1 Univ. Grenoble Alpes, CNRS, CEA, Institut de Biologie Structurale (IBS), 71, Avenue des martyrs, 38044 Grenoble, France

2 NMR-Bio, 5 place Robert Schuman, 38025 Grenoble, France

3 Discovery Technologies, Merck KGaA, Frankfurter Straße 250, 64293 Darmstadt, Germany

4 Univ. Grenoble Alpes, CEA, CNRS, IRIG, CBM, 38000 Grenoble, France

et al. 2012; Ayala et al. 2020). On the other hand, as leucine, valine and isoleucine residues are at the end of irreversible metabolic pathways in *E. coli*, precursors can be incorporated in the growth medium for their cost-effective labelling. The first precursors introduced to label isoleucine or leucine and valine residues were 2-keto acids: α-ketobutyrate (Gardner et al. 1997) and α-ketoisovalerate (Goto et al. 1999; Hajduk et al. 2000; Gross et al. 2003), respectively. However, α-ketoisovalerate used as a precursor for leucine and valine residues is leading to a non-stereospecific labelling of both pro-*S* and pro-*R* $^{13}C^{1}H_3$ groups resulting in overcrowded spectra for high molecular weight proteins, even when both sensitivity and resolution have been improved using a non-stereospecific $^{13}C^{1}H_3/^{12}C^{2}H_3$ α-ketoisovalerate (Tugarinov and Kay 2004b).

To prevent peak overlaps and to facilitate studies of large molecular weight assemblies, methyl labelled acetolactate has been used as an alternative precursor (Gans et al. 2010). This latter enables the stereospecific $^{13}C^{1}H_3$-labelling of valine and leucine methyl groups and therefore halves the number of peaks observed whilst improving by two the sensitivity of the spectrum as compared to labelling at 50% pro-*S* and 50% pro-*R* methyl moieties using optimized $^{13}C^{1}H_3/^{12}C^{2}H_3$ α-ketoisovalerate (Tugarinov and Kay 2004b). This interesting precursor also enhances the intensity of NOE cross peaks and increases the distance threshold at which NOE cross peaks can be detected by 20% (Gans et al. 2010).

However, despite the tremendous progress made in protocols to selectively introduce protonated methyl groups in perdeuterated proteins, sequence specific assignment, essential for analyzing a variety of NMR data, remains an important challenge for large molecular assemblies. Several methods to solve the bottleneck of assignment of large proteins have been developed including parallel mutagenesis strategies (Amero et al. 2011), structure based approaches using the analysis of NOE cross peaks with various programs (Pritišanac et al. 2020), MAP-XSII (Xu and Matthews 2013), FLAMEnGO 2.0 (Chao et al. 2014), MAGMA (Pritišanac et al. 2017), MAGIC (Monneau et al. 2017), MethylFLYA (Pritišanac et al. 2019), MAUS (Nerli et al. 2021) or the "divide-and-conquer" approach (Gelis et al. 2007; Sprangers and Kay 2007) which is based on separation of large proteins into smaller fragments, assigning these and transferring the assignment back to the full-length protein. For proteins of moderate molecular weight or fragments of large assemblies for which backbone assignment is available, it is possible to connect methyl resonances to those of the backbone. This requires a sample with $^{13}C^{1}H_3$ labelled methyl groups connected to the backbone by a linear ^{13}C-chain. With such a sample, transfer from the assigned backbone to the methyl groups can be achieved using either unidirectional transfer from methyl groups to HN using (HM)CM(CGCBCA)NH experiments (Tugarinov and Kay 2003) or 'out and back' HCC relay triple resonance experiments (Tugarinov and Kay 2003; Ayala et al. 2012; Mas et al. 2013).

A combination of these techniques to assign methyl groups in addition with a stereospecific labelling, enhancing the sensitivity of the spectrum by a factor two and significantly reducing signal overlap, should lead to straightforward leucine and valine methyl group assignment. Labelling schemes, connecting non-stereospecifically both leucine and valine methyl groups (Tugarinov and Kay 2003), or only pro-*R* methyl moieties (Mas et al. 2013; Kerfah et al. 2015a), to the assigned backbone are already available. However, with such precursors additional samples are required either to stereospecifically assign the methyl group (Tugarinov and Kay 2004a; Gans et al. 2010) or to link the pro-*S* methyl to the pro-*R* one (Mas et al. 2013; Kerfah et al. 2015a). Here we introduce the synthesis of a new dissymmetric $^{13}C/^{2}H$-labelled acetolactate, a precursor that allows to directly connect the assigned leucine and valine backbone atoms to the pro-*S* methyl groups via a linear ^{13}C chain using only one sample. This new labelling scheme has been applied to the N-terminal domain of human HSP90 (HSP90-NTD) and we present here the full assignment of methyl moieties of this protein.

Materials and methods

Synthesis of 1,2,3-[$^{13}C_3$]-4,4,4-[$^{2}H_3$]-acetolactate

Synthesis of ethyl 1,2,3-[$^{13}C_3$]-3-oxo-butanoate

A solution of LiHMDS (7.80 g, 46.6 mmol, 2.1 equiv.) in freshly dried THF (150 mL) was cooled to − 78 °C under argon. 1,2-[$^{13}C_2$]-ethyl acetate (2.00 g, 22.2 mmol, Cambridge Isotope Laboratory, CIL) was added dropwise. The resulting solution was stirred at − 78 °C for 15 min, then 1-[^{13}C]-acetyl chloride (1.60 mL, 22.2 mmol, 1 equiv., CIL) was added dropwise. The resulting mixture was stirred at − 78 °C for additional 30 min then quenched by addition of a 20% aqueous solution of ^{1}HCl (15 mL). After three extractions with Et_2O, the organics were combined, washed with saturated $Na^{1}HCO_3$ solution then dried over Na_2SO_4. Concentration under vacuum affords the desired product (2.88 g) which was used in the next step without further purification).

^{1}H NMR:($C^{2}HCl_3$), δ: 4.21 (dq, O-CH_2, $^{3}J(^{1}H$-$^{1}H) = 7.1$ Hz, $^{3}J(^{1}H$-$^{13}C) = 3.2$ Hz, 2H); 3.46 (dt, ^{13}C-$^{13}CH_2$-^{13}C, $^{2}J(^{1}H$-$^{13}C) = 6.5$ Hz, $^{1}J(^{1}H$-$^{13}C) = 130.1$ Hz, 2 H); 2.28 (dd, CH_3-^{13}C, $^{3}J(^{1}H$-$^{13}C) = 1.4$ Hz, $^{2}J(^{1}H$-$^{13}C) = 6.1$ Hz, 3H), 1.30 (t, OCH2–CH3, $^{3}J(^{1}H$-$^{1}H) = 7.1$ Hz, 3H).

Synthesis of ethyl 1,2,3-[$^{13}C_3$]-2-[$^{13}C^{1}H_3$]-3-oxo-butanoate $^{13}C^{1}H_3$-I (752 mL, 11.99 µmoles, 1.1 equiv, CIL)

was slowly added to a solution of ethyl 1,2,3-[$^{13}C_3$]-3-oxo butanoate (1.45 g, 10.90 mmol) in EtO^1H (50 mL) cooled to 0 °C before addition of K_2CO_3 (1.66 g, 11.99 mmol, 1.1 equiv.). The resulting suspension was warmed to room temperature then stirred for 18 h. The mixture was concentrated to the fifth before addition of a large volume of Et_2O. Excess of K_2CO_3 was filtered off and the filtrate concentrated under vacuum to the fifth before a further addition of Et_2O and a second filtration. Concentration under vacuum affords the desired product (1.03 g) as a colorless oil which was used in the next step without further purification.

^{1}H NMR: (C^2HCl_3), δ: 4.21 (dq, O-CH_2, ^{3}J(^{1}H-^{1}H) = 7.1 Hz, ^{3}J(^{1}H-^{13}C) = 3.0 Hz, 2H); 3.50 (dm, ^{13}C-^{13}CH-^{13}C, ^{1}J(^{1}H-^{13}C) = 129.0 Hz, 1H); 2.24 (dd, CH_3-^{13}C, ^{3}J(^{1}H-^{13}C) = 1.3 Hz, ^{2} J(^{1}H-^{13}C) = 6.0 Hz, 3H), 1.36 (dm, $^{13}CH_3$, ^{1}J(^{1}H-^{13}C) = 129.0 Hz, 3H), 1.28 (t, OCH_2-CH_3, ^{3}J(^{1}H-^{1}H) = 7.1 Hz, 3H).

Synthesis of ethyl 1,2,3-[$^{13}C_3$]-2-[$^{13}C^1H_3$]-2-[O^1H]-3-oxo-butanoate To a solution of ethyl 1,2,3-[$^{13}C_3$]-2-[$^{13}C^1H_3$]-3-oxo-butanoate (995 mg, 6.72 mmol) in DMSO (8 mL), Cs_2CO_3 was added (440 mg, 1.35 mmol, 0.2 equiv.). After O_2 bubbling for 15 min, P(OEt)$_3$ (233 mL, 1.35 mmol, 0.2 equiv.) was added. The resulting solution was stirred under O_2 atmosphere for 20 h. A large volume of Et_2O was then added followed by a saturated NaCl solution. The resulting phases were separated and the aqueous one was extracted one more time with Et_2O. The organics were combined, dried over Na_2SO_4 then concentrated under vacuum to obtain ethyl 1,2,3-[$^{13}C_3$], 2-[$^{13}C^1H_3$], 2-[O^1H]-3-oxo-butanoate as a yellow oil (1.142 g) and pure enough to be used in the next step without further purification.

^{1}H NMR:(C^2HCl_3), δ: 4.25 (dq, O-CH_2, ^{3}J(^{1}H-^{1}H) = 7.1 Hz, ^{3}J(^{1}H-^{13}C) = 3.2 Hz, 2H); 4.17–4.24 (m, OH, 1H), 2.27 (dd, CH_3-^{13}C, ^{3}J(^{1}H-^{13}C) = 1.1 Hz, ^{2}J(^{1}H-^{13}C) = 6.1 Hz, 3H), 1.36 (dm, $^{13}CH_3$, ^{1}J(^{1}H-^{13}C) = 134.2 Hz, 3H), 1.28 (t, OCH2–CH3, ^{3}J(^{1}H-^{1}H) = 7.1 Hz, 3H).

Synthesis of sodium 1,2,3-[$^{13}C_3$]-2-[$^{13}C^1H_3$]-2-[O^2H]-3-oxo-4,4,4-[$^2H_3$]-butanoate To a solution of ethyl 1,2,3-[$^{13}C_3$]-2-[$^{13}C^1H_3$]-2-[O^1H]-3-oxobutanoate (1.09 g) in ^2H_2O (4 mL), 0.4 equivalents of a solution of NaO^2H (2.5 M) in ^2H_2O were added dropwise in 40 min, using a syringe pump under argon. As soon as the addition was completed, ^{1}H NMR was carried out on a sample (few µL) in ^2H_2O in order to calculate the conversion percentage [*ratio* between the amount of hydrolyzed product (quadruplet at 1.60 ppm) and the amount of starting material (quadruplet at 1.7 ppm)]. 1.1 equivalents of a NaO^2H solution (2.5 M) was added over 30 min with the syringe pump. As soon as the addition was completed, an extraction with diethyl ether was carried out in order to remove the by-product coming from the previous step whose NMR signals prevent a good follow-up of the hydrogen/deuterium (^{1}H/^{2}H) exchange on the 4-CH_3 (2.2 ppm). The ^{1}H/^{2}H exchange on 4-CH_3 was then monitored by ^{1}H NMR and carried out by successive addition of NaO^2H (2.5 M) until the integral of the doublet corresponding to the CH_3 reaches the value of 0.1 when the quadruplet at 1.6 ppm integrates for 1.5. The reaction was immediately neutralized with a concentrated ^{2}HCl solution to neutral pH and then buffered with Tris–^{1}HCl, (1.0 M, pH 7.5 in ^2H_2O). The concentration of the resulting solution was then determined by ^{1}H NMR using methanol or acetonitrile as internal reference. The final product (3.02 mmol) was stored at − 80 °C.

^{1}H NMR:(C^2HCl_3), δ: 1.37 (dq, $^{13}CH_3$, ^{1} J(^{1}H-^{13}C) = 129.3 Hz, ^{2} J(^{1}H-^{13}C) = 3.9 Hz, 1 H).

Preparation of isotopically labelled HSP90-NTD samples

Escherichia coli BL21-DE3-RIL cells transformed with a pET-28 plasmid encoding the N-Terminal domain of HSP90 α from *Homo Sapiens* (HSP90-NTD) with a His-Tag and a TEV cleavage site were progressively adapted in three stages over 24 h to M9/^2H_2O. In the final culture, bacteria were grown at 37 °C in M9 medium with 99.85% ^2H_2O (Eurisotop), 1 g/L $^{15}N^1H_4Cl$ (Sigma Aldrich) and 2 g/L D-glucose-d_7 (for U-[^{2}H, ^{12}C, ^{15}N] HSP90-NTD samples) or D-glucose-$^{13}C_6$-d_7 (CIL) (for U-[^{2}H, ^{13}C, ^{15}N] HSP90-NTD samples).

For methyl specifically labelled samples, the methyl labelled precursors or amino-acids were added to the media when the O.D at 600 nm reached 0.6 (Kerfah et al. 2015c):

- Labelling scheme (A): for production of the U-[^{2}H, ^{15}N, ^{13}C], Ile-[2,3,4,4-$^2H_4$; 1,2,3,4-$^{13}C_4$; $^{13}C^1H_3]^{\delta 1}$/[$^{12}C^2H_3]^{\gamma 2}$], Leu-[2,3,3,4-$^2H_4$; 1,2,3,4-$^{13}C_4$; [$^{13}C^1H_3]^{pro\text{-}S}$/[$^{12}C^2H_3]^{pro\text{-}R}$], Val-[2, 3-$^2H_2$; 1,2,3-$^{13}C_3$; [$^{13}C^1H_3]^{pro\text{-}S}$/[$^{12}C^2H_3]^{pro\text{-}R}$] HSP90-NTD, a solution containing the sodium 1,2,3-[$^{13}C_3$]-2-[$^{13}C^1H_3$]-2-[O^2H]-3-oxo-4,4,4-[$^2H_3$]-butanoate precursor was added at a concentration of 172 mg/L 1 h before induction. 40 min later (20 minutes before induction) a solution containing 60 mg/L of sodium (*S*)-2-hydroxy-2-(1′,1′-[$^2H_2$], 1′,2′-[$^{13}C_2$])ethyl-3-oxo-1,2,3-[$^{13}C_3$]-4,4,4-[$^2H_3$]butanoate (Kerfah et al. 2015a) was added to the medium.
- Labelling scheme (B): for production of the U-[^{2}H, ^{15}N, ^{12}C], Ala-[$^{13}C^1H_3]^{\beta}$, Met-[$^{13}C^1H_3]^{\varepsilon}$, Leu/Val-[$^{13}C^1H_3]^{pro\text{-}S}$, Ile-[$^{13}C^1H_3]^{\delta 1}$, Thr-[$^{13}C^1H_3]^{\gamma}$ HSP90-NTD, a HLAM-$A^{\beta}I^{\delta 1}M^{\varepsilon}LV^{proS}T^{\gamma}$ kit, purchased from NMR-Bio, was added before induction according to the manufacturer's protocol.
- Labelling scheme (C): U-[^{2}H, ^{15}N, ^{12}C] samples labelled on a single type of methyl group were produced in small scales (21 mL) to identify A^{β}, M^{ε}, T^{γ} methyl type (3 samples) or to complete assignment using single point mutants (Amero et al. 2011) (33 samples, list of mutants

presented in the legend of Fig. S4). Single point amino acid mutations were generated by GeneCust. For each of these samples a single type of methyl groups was labelled by addition of the corresponding NMR-Bio kit (SLAM-A^{β}, SLAM-M^{ε}, SLAM-$I^{\delta 1}$, SLAM-T^{γ} or DLAM-LV^{proS}) in M9/^2H_2O media 1 h before induction.

Protein production was induced by the addition of IPTG to a final concentration of 0.5 mM. The cultures were grown overnight at 20 °C before harvesting. Cells were collected by centrifugation at 5500 g for 20 min at 4 °C then lysed by sonication on ice in a buffer containing 20 mM phosphate sodium buffer at pH 7.4, 0.5 M NaCl, 0.05% β-ME, antiprotease (cOmplete® EDTA free, 1 tablet for 50 mL), 50 µg/mL DNAse (Sigma Aldrich), 50 µg/mL RNAse (Euromedex), and 0.25 mg/mL Lysozyme (Euromedex). After removal of cell debris by centrifugation (45,000×g, 30 min, 4 °C), the supernatant was purified using an affinity chromatography step (Ni-NTA, Superflow, QIAGEN) (labelling scheme A, B and C), followed by a size exclusion chromatography step (16/600 Superdex 75 PG, GE Healthcare) (labelling schemes A and B only). The gel filtration column was run with an isocratic step of the NMR buffer (20 mM Hepes, 150 mM NaCl, 1 mM TCEP, pH 7.5).

The HSP90-NTD proteins were concentrated, using an Amicon® 4 Centrifugal Filter Unit with a 10,000 MWCO (Merck), either in a 90%/10% $^1H_2O/^2H_2O$ or in a 100% ^2H_2O buffer containing 20 mM Hepes, 150 mM NaCl, 1 mM TCEP, pH 7.5. For labelling schemes A and B, samples were concentrated to 0.5 mM and 200 µL of each sample was loaded in 4 mm shigemi tube. The wild type and single point mutants of HSP90-NTD proteins labelled on only one methyl type (labelling scheme C) were concentrated at [0.1–0.4] mM and 40 µL of each sample was loaded in a 1.7 mm NMR tube.

NMR Spectroscopy

All NMR experiments acquired on HSP90-NTD samples were recorded at 298 K. 2D ^1H-^{13}C SOFAST methyl TROSY (Amero et al. 2009) experiments to identify each methyl type as well as to assign individual methyl signals using single point mutants were recorded for an average duration of ~ 1.5 h each, on a spectrometer operating at a ^1H frequency of 850 MHz and equipped with a 1.7 mm cryogenically cooled, pulsed-field-gradient triple-resonance probe. All other NMR experiments were acquired using Bruker Avance III HD spectrometers equipped with 5 mm cryogenic probes (operating at a ^1H frequency of 600 or 950 MHz).

The 3D HCC, HC(C)C and HC(CC)C experiments (Tugarinov and Kay 2003; Ayala et al. 2009, 2012; Mas et al. 2013) were acquired on a spectrometer operating at a ^1H frequency of 600 MHz for a total duration of 4 days, using a 0.5 mM sample of U-[^2H, ^{15}N, ^{13}C], Ile-[2,3,4,4-$^2H_4$; 1,2,3,4-$^{13}C_4$; $[^{13}C^1H_3]^{\delta 1}/[^{12}C^2H_3]^{\gamma 2}$], Leu-[2,3,3,4-$^2H_4$; 1,2,3,4-$^{13}C_4$; $[^{13}C^1H_3]^{pro\text{-}S}/[^{12}C^2H_3]^{pro\text{-}R}$], Val-[2,3-$^2H_2$; 1,2,3-$^{13}C_3$; $[^{13}C^1H_3]^{pro\text{-}S}/[^{12}C^2H_3]^{pro\text{-}R}$] HSP90-NTD. The interscan delay was adjusted to 0.5–0.6 s, the heteronuclear ^1H -> ^{13}C transfer delay was set to 4 ms ($1/(2\times{}^1J_{HC})$) and the homonuclear ^{13}C -> ^{13}C transfer delay was fixed to 12.5 ms. The acquisition times were adjusted to 8-10.7 ms in the ^{13}C indirect dimension, and to 70 ms in ^1H direct dimension.

For the sequential assignment of backbone resonances, a set of 6 BEST-TROSY 3D triple resonance experiments (HNCA, HN(CA)CB, HNCO, HN(CA)CO, HN(CO)CA and HN(COCA)CB (Favier and Brutscher 2019) were acquired on a Bruker Avance III HD spectrometer equipped with a cryogenic probe and operating at a ^1H frequency of 600 MHz for a total duration of 11 days using a 0.5 mM sample of the U-[^2H, ^{15}N, ^{13}C] HSP90-NTD.

The 3D CCH HMQC-NOESY-HMQC NMR experiment (Tugarinov et al. 2005; Törner et al. 2020) was recorded over 3 days on a spectrometer operating at a ^1H frequency of 950 MHz using a 0.5 mM sample of U-[^2H, ^{15}N, ^{12}C], Ala-$[^{13}C^1H_3]^{\beta}$, Met-$[^{13}C^1H_3]^{\varepsilon}$, Leu/Val-$[^{13}C^1H_3]^{pro\text{-}S}$, Ile-$[^{13}C^1H_3]^{\delta 1}$, Thr-$[^{13}C^1H_3]^{\gamma}$ HSP90-NTD. The interscan delay was set to 1.1 s. The heteronuclear ^1H -> ^{13}C transfer delay was set to 4 ms ($1/(2\times{}^1J_{HC})$). The acquisition times in the ^{13}C indirect dimension were set to 24.6 ms, (t_{1max}) and to 18.6 ms (t_{2max}). In the ^1H direct dimension t_{3max} was fixed to 80 ms. The NOE mixing period was set to 500 ms to detect a maximum number of long-range intermethyl NOEs.

Data processing and analysis

All data were processed and analyzed using nmrPipe/nmrDraw (Delaglio et al. 1995) and CcpNMR (Vranken et al. 2005). Automated methyl assignment was performed using MAGIC software (Monneau et al. 2017) using the reference structure of HSP90-NTD (PDB: 1YES). Input NOE lists for MAGIC were created with CcpNMR. MAGIC was run with a score threshold factor of 1 and distance thresholds of 7–10 Å using all inter methyl NOEs detected (S/N threshold of 5 was used) and given the assignment of isoleucines, leucines and valines previously obtained by the three 'out and back' HCC experiments as well as alanine, methionine and threonine methyl groups assigned by mutagenesis as additional input.

Results and discussion

To assign methyl groups of large perdeuterated proteins, previously assigned backbone resonances of these proteins can be used (Tugarinov and Kay 2003; Ayala et al. 2012; Mas et al. 2013). Nonetheless, to do so, methyl groups need to be

connected, via a linear chain of ^{13}C, to the backbone atoms in order to be able to apply optimized experiments to high molecular weight proteins. Strategies to label stereospecifically leucine and valine pro-*R* methyl groups and to connect them to backbone nuclei have already been proposed (Mas et al. 2013). However, it has to be noted that pro-*S* methyl groups are often chosen over pro-*R* methyl groups as they are both easier and cheaper to label stereospecifically (Gans et al. 2010). In order to simplify assignment of pro-*S* methyl groups using already assigned backbone resonances and to avoid the need of an additional sample to link the pro-*S* methyl to the pro-*R* one, a sample connecting the pro-*S* $^{13}C^1H_3$-methyl groups to the backbone atoms by a linear ^{13}C-chain and labelled with $^{12}C^2H_3$ on pro-*R* methyl moieties to avoid signal loss would be optimal.

Synthesis of optimally labelled acetolactate precursors and proteins

Taking into account the specificity of leucine/valine metabolic pathway in *E. coli*, such an optimal labelling scheme can be achieved in M9/^2H_2O medium using $^{13}C/^2H$ glucose (Kerfah et al. 2015c) as a carbon source together with 1,2,3-[$^{13}C_3$]-2-[$^{13}C^1H_3$]-2-[O^2H]-3-oxo-4,4,4-[$^2H_3$]-butanoate as suitably labelled acetolactate precursor. However, this latter cannot be synthetized from commercially available materials by the traditional route starting from acetoacetate (Gans et al. 2010) since the corresponding labelled starting material is not commercially available. Indeed, acetolactate chemical synthesis is achieved by reaction of iodomethane on acetoacetate (Gans et al. 2010). Whilst both $^{13}C^1H_3$ labelling of the methyl substituent in position 2 and deuteration of the methyl group in position 4 can be obtained using $^{13}C^1H_3I$ as a starting synthesis material and hydrogen/deuterium exchange in controlled basic conditions (Gans et al. 2010), respectively, the ^{13}C labelling of only the first three carbons of the main chain using commercially available labelled acetoacetate materials is not achievable. Therefore, as acetoacetate can be obtained by condensation of two acetate moieties (Epstein et al. 1977), we decided to set up a synthesis of dissymmetrically labelled acetoacetate starting from commercially available ethyl 1, 2-[$^{13}C_2$]- acetate with 1-[^{13}C]-acetyl chloride. Based on reported procedures, we established a 4-step synthesis (Fig. 1) allowing to prepare the desired precursor with an overall yield of 27%. In brief, the dissymmetry is achieved by the Claisen condensation of ethyl-1,2,[$^{13}C_2$]-acetate with 1-[^{13}C]-acetyl chloride using an optimization of a reported procedure (Epstein et al. 1977) (a), followed by an alkylation in position 2 using $^{13}C^1H_3I$ (b) and a subsequent hydroxylation in position 2 (c). Finally, the last step combining both saponification of the ester and a hydrogen/deuterium exchange in position 4 is performed under controlled basic conditions (d). This last step is very delicate and requires a fine control of the basic condition as a methyl rearrangement above a pH of 13.5 can take place resulting in the interconversion of both methyl groups (Gans et al. 2010). Steps (b) and (c) are not stereoselective, hence, products of these latter steps were produced as racemic mixtures. The optimally labelled acetolactate, unstable at room temperature, was aliquoted and stored at -80 °C.

The synthetized acetolactate precursor was incorporated in *E. coli*. M9/^2H_2O culture media without any further purification steps to label the overexpressed protein. Frozen acetolactate vials were thawed right before

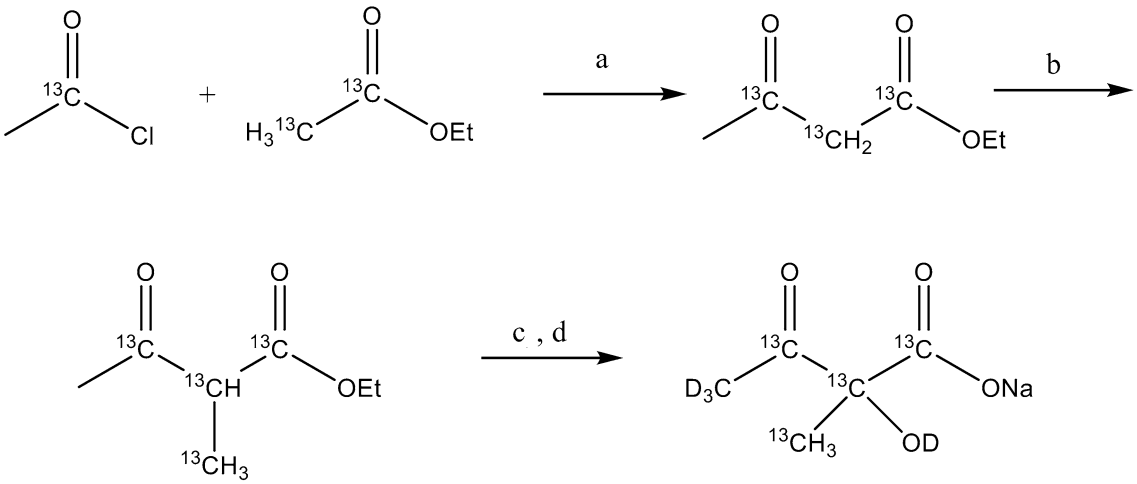
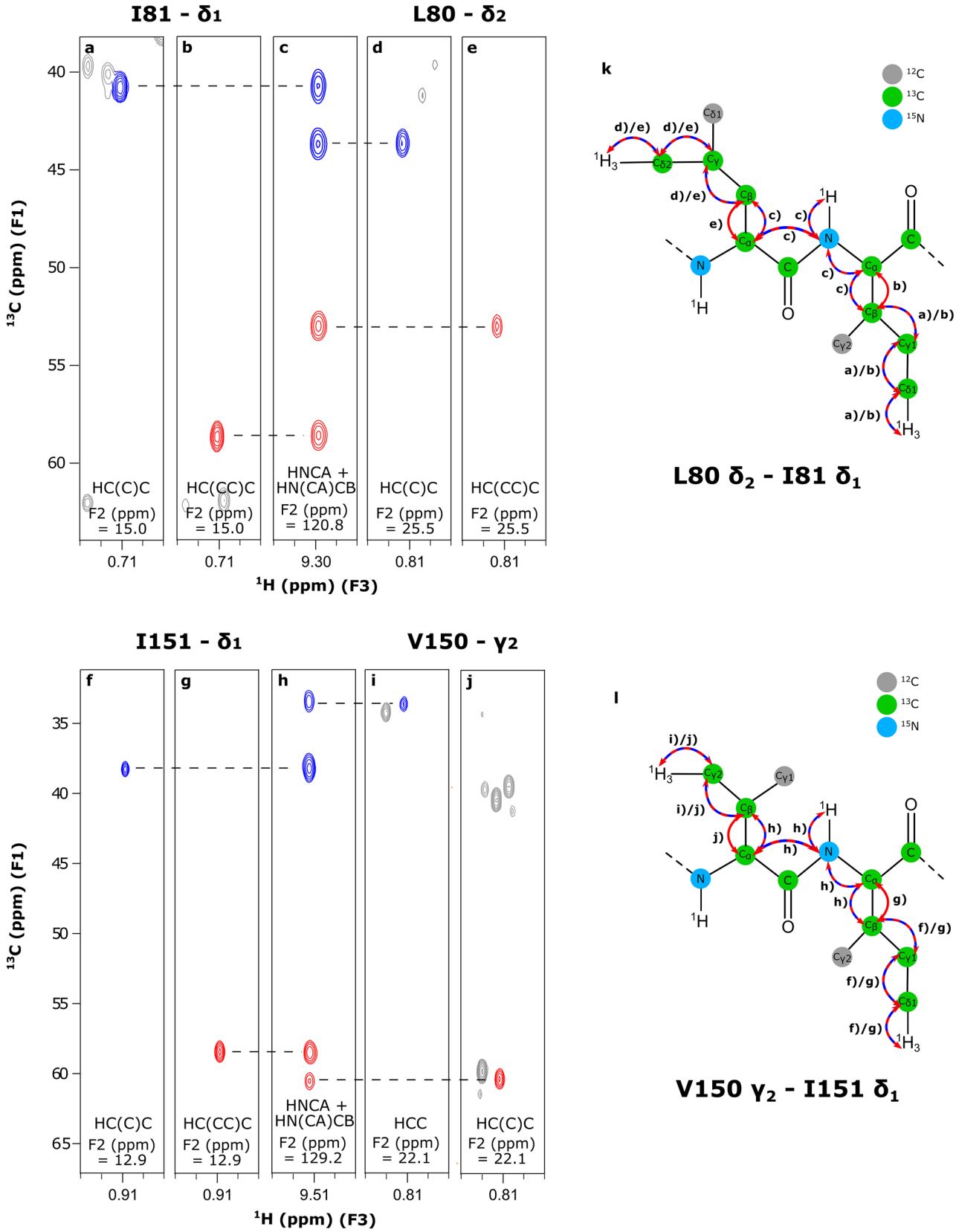


Fig. 1 Synthetic scheme for preparation of specifically labelled sodium 1,2,3-[$^{13}C_3$]-2-[$^{13}C^1H_3$]-2-[O^2H]-3-oxo-4,4,4-[$^2H_3$]-butanoate. **a** i—LiHMDS (2.1 equiv.), dry THF, − 78 °C; ii—1,2-[$^{13}C_2$]-ethyl acetate (1 equiv.), − 78 °C, 15 min.; iii—1-[^{13}C]-acetyl chloride (1.0 equiv.), − 78 °C, 30 min.; iv—^{1}HCl 20%; **b** K_2CO_3 (1.1 equiv.), $^{13}C^1H_3$-I (1.1 equiv.), 0 °C, 18 h, EtO^1H; **c** Cs_2CO_3 (0.2 equiv.), $P(OEt)_3$ (0.2 equiv.), O_2, DMSO, 20 h. **d** i—NaO^2H (2.5 M), ^2H_2O; ii—^{2}HCl 35%; Tris buffer pH 7.5; 27% overall yield


I81 - δ1
L80 - δ2
a
b
c
d
e
40
45
50
55
60
13C (ppm) (F1)
HC(C)C
F2 (ppm)
= 15.0
HC(CC)C
F2 (ppm)
= 15.0
HNCA +
HN(CA)CB
F2 (ppm)
= 120.8
HC(C)C
F2 (ppm)
= 25.5
HC(CC)C
F2 (ppm)
= 25.5
0.71
0.71
9.30
0.81
0.81
1H (ppm) (F3)
k
12C
13C
15N
L80 δ2 - I81 δ1
I151 - δ1
V150 - γ2
f
g
h
i
j
35
40
45
50
55
60
65
13C (ppm) (F1)
HC(C)C
F2 (ppm)
= 12.9
HC(CC)C
F2 (ppm)
= 12.9
HNCA +
HN(CA)CB
F2 (ppm)
= 129.2
HCC
F2 (ppm)
= 22.1
HC(C)C
F2 (ppm)
= 22.1
0.91
0.91
9.51
0.81
0.81
1H (ppm) (F3)
l
12C
13C
15N
V150 γ2 - I151 δ1

◂**Fig. 2** Assignment transfer from the backbone to Leu$^{pro\text{-}S}$, Val$^{pro\text{-}S}$ and Ile-δ_1 methyl groups of HSP90-NTD. Examples of 2D-extracts from 3D 'out and back' HCC (i), HC(C)C (**a**, **d**, **f** and **j**) and HC(CC)C (**b**, **e** and **g**) experiments correlating ^{1}H (F_3) and ^{13}C (F_2) methyl resonances with $^{13}C_\beta$ (blue) or $^{13}C_\alpha$ (red) in F_1 dimension. Panels c and h display the corresponding 2D HNCA and HN(CA)CB extracts for Ile-81 (c) and Ile-151 (**h**) allowing to connect Leu-80-δ_2 (pro-*S*), Ile-81-δ_1, Val-150-γ_2 (pro-*S*), Ile-151-δ_1 methyl groups to previously assigned backbone atoms. 3D spectra were recorded on an NMR spectrometer operating at a proton frequency of 600 MHz using the U-[^{2}H, ^{15}N, ^{13}C], Ile-[2,3,4,4-$^{2}H_4$; 1,2,3, 4-$^{13}C_4$; $^{13}C^{1}H_3]^{\delta1}$/$[^{12}C^{2}H_3]^{\gamma2}$], Leu-[2,3,3,4-$^{2}H_4$; 1,2,3,4-$^{13}C_4$; $[^{13}C^{1}H_3]^{pro\text{-}S}$/$[^{12}C^{2}H_3]^{pro\text{-}R}$], Val-[2,3-$^{2}H_2$; 1,2,3-$^{13}C_3$; $[^{13}C^{1}H_3]^{pro\text{-}S}$/$[^{12}C^{2}H_3]^{pro\text{-}R}$] labelled sample or U-[$^{2}H$, ^{15}N, ^{13}C] labelled sample (3D HNCA and HN(CA)CB). **k** and **l** represent magnetization transfer schemes correlating with the strips (**a**, **b**, **c**, **d**, **e**) and (**f**, **g**, **h**, **i**, **j**), respectively. Except when specified, all hydrogen atoms are ^{2}H

addition in the culture medium to avoid degradation or methyl rearrangement. It has to be noted that only the 2-(*S*) stereoisomer of acetolactate is converted in vivo by ketol–acid reductoisomerase (EC1.1.1.86) and dihydroxy-acid dehydratase (EC 4.2.1.9) to form the stereospecifically labelled 2-keto-isovalerate. This latter is afterwards directly converted into valine or combined with $^{13}C/^{2}H$-pyruvate, derived from the $^{13}C/^{2}H$-glucose, to produce leucine with the desired labelling pattern. The 2-(*R*) stereoisomer of acetolactate is, itself, not a substrate of ketol–acid reductoisomerase and hence, induces no scrambling (Gans et al. 2010). The synthetized 1,2,3-[$^{13}C_3$]-2-[$^{13}C^{1}H_3$]-2-[$O^{2}H$]-3-oxo-4,4,4-[$^{2}H_3$]-butanoate can be mixed with other known precursors allowing also to connect methyl groups, such as Ile-δ_1 (Kerfah et al. 2015a, b; Törner et al. 2020), Ile-γ_2 (Ayala et al. 2012) or Ala-β (Ayala et al. 2009; Kerfah et al. 2015a; Törner et al. 2020), to backbone nuclei using a linear ^{13}C chain. In this study, we chose to label our sample on both leucine/valine pro-*S* and isoleucine-δ_1 methyl moieties, by adding Ile-δ_1 precursor (sodium (S)-2-hydroxy-2-(1′,1′-[$^{2}H_2$], 1′, 2′-[$^{13}C_2$])ethyl-3-oxo-1,2,3-[$^{13}C_3$]-4,4,4-[$^{2}H_3$]-butanoate – (Kerfah et al. 2015a, b) together with our new optimized acetolactate. The isoleucine precursor was added in the culture medium 40 min later than the new precursor in order to take into account co-incorporation incompatibilities between both the leucine/valine precursor and the isoleucine one. Indeed, enzymes from ILV-pathway have a tendency to process preferentially isoleucine precursor instead of leucine/valine precursors (Kerfah et al. 2015b, c). Incorporation of both precursors during the protein expression did not lead to a significantly different HSP90-NTD yield with regards to the yields obtained in standard M9/$^{2}H_2O$ media. No scrambling was detected neither to the pro-*R* methyls groups of leucine and valine nor to the isoleucine-γ_2 site (Fig. S1a).

Connection of $I^{\delta1}L^{\delta2}V^{\gamma2}$ methyl groups to C_α and C_β atoms

Using this new precursor, it is possible to directly correlate pro-*S* methyl groups of leucine and valine residues to their respective C_α and C_β. We decided to apply this strategy to the N-terminal domain of human HSP90 (HSP90-NTD), an extensively studied protein, whose isoform assignment (α and β), including partial assignment of its methyl groups, is available (Jacobs et al. 2006; Elif Karagöz et al. 2011; Park et al. 2011; Lescanne et al. 2017, 2018). Here, we have focused on the N-terminal domain of the α isoform, a 29 kDa protein that contains 20 isoleucine, 18 leucine and 11 valine residues. This dynamic protein is particularly challenging to assign using automatic methyl assignment methods and reported success rates are ranging from 27 to 69% (Pritišanac et al. 2017, 2019, 2020; Monneau et al. 2017). In our hands, only 34% of the methyl groups (30/87) could be assigned automatically with (1) a single assignment, (2) a high NOE assignment completeness of the strip related to each peak (> 50%) and (3) a high total confidence score value (≥ 7) (Table S1) using the HSP90-NTD X-ray structure (PDB: 1YES) and experimentally detected NOE network. Therefore, this protein is a good candidate to assess our experimental strategy based on new precursors.

To do so, 200 µL at 0.5 mM of an optimally labelled sample was used to acquire three 'out and back' HCC, HC(C)C and HC(CC)C experiments (Tugarinov and Kay 2003; Ayala et al. 2009, 2012; Mas et al. 2013) connecting labelled $I^{\delta1}L^{\delta2}V^{\gamma2}$ methyls groups to $I^{\gamma1}L^{\gamma}V^{\beta}$, $I^{\beta}L^{\beta}V^{\alpha}$ and $I^{\alpha}L^{\alpha}$ resonances. 100 and 94% of the expected C_β and C_α coherences, respectively, were observed for the 29 kDa HSP90-NTD at 298 K (τ_C ca. 20 ns) (Fig. 2). The three missing C_α resonances correspond to residues L56, I26 and I110, two of them being affected by extensive line broadening due to conformational exchange. With such a high percentage of observed C_α and C_β resonances we demonstrate the applicability of this method for medium size proteins. In order to validate the strategy for larger proteins, the labelling scheme was applied to the 87 kDa hetero hexameric protein prefoldin from Pyrococcus horikoshii, containing 2α and 4β-subunits. Only the β-subunits were labelled with the optimal labelling schemes described above, whilst the α-subunits remained perdeuterated (Fig. S1b). Three HCC experiments were collected on this 0.2 mM sample of prefoldin at 310 K (τ_C ca. 60 ns). 95 and 60% of the expected C_β and C_α coherences, respectively, were observed (Fig. S2) despite the high molecular weight of the protein and the presence of doubled peaks due to the presence of two inequivalent β-subunits in the $\alpha_2\beta4$ hexameric prefoldin (Ohtaki et al. 2008). Such HCC experiments were also acquired at 343 K (τ_C ca. 30 ns) on this hyperthermophilic prefoldin sample, enabling transfer

of assignment between backbone atoms and methyl groups (Törner et al. 2021).

Application to the sequence specific assignment of $I^{\delta1}L^{\delta2}V^{\gamma2}$ methyl groups of HSP90-NTD

Backbone sequential assignment was performed using 6 'BEST-TROSY' triple resonance experiments (Favier and Brutscher 2019). C_α and C_β resonances were assigned for 89% and 80% of the residues of HSP90-NTD respectively, excluding the loosely structured N-terminal [1–16] and C-terminal [225–236] regions. The segment L103-T115, that covers the ligand binding site, is invisible by NMR due to dynamics in the µs-ms timescale. Transfer of sequentially assigned backbones resonances to isoleucine-δ_1, leucine-δ_2 and valine-γ_2 methyl groups was achieved using 'out and back' HCC experiments acquired on a U-[^{2}H, ^{15}N, ^{13}C], Ile-[2,3,4,4-$^2H_4$; 1,2,3,4-$^{13}C_4$; $[^{13}C^1H_3]^{\delta1}/[^{12}C^2H_3]^{\gamma2}$], Leu-[2,3,3,4-$^2H_4$; 1,2,3,4-$^{13}C_4$; $[^{13}C^1H_3]^{pro\text{-}S}/[^{12}C^2H_3]^{pro\text{-}R}$], Val-[2,3-$^2H_2$; 1,2,3-$^{13}C_3$; $[^{13}C^1H_3]^{pro\text{-}S}/[^{12}C^2H_3]^{pro\text{-}R}$] labelled HSP90-NTD sample. 2D extracts from the HCC experiments were compared with the corresponding ones from the 3D HNCA and HN(CA)CB experiments and using C_α and C_β resonances, all the $I^{\delta1}L^{\delta2}V^{\gamma2}$ methyl groups could be unambiguously connected to previously assigned backbone atoms (Fig. 2). The assignment was transferred in one step, very simply and efficiently without having to correct for the isotopic shifts (Kerfah et al. 2015a). One must note that the methyl-δ_1 of Ile-33 and Ile-128 are superimposed in the 2D methyl-TROSY spectrum, but were unambiguously connected to the C_α resonances of both amino acids (Fig. S3). Remained unassigned only methyl groups of L103, I104, L107 and I110 for which backbone atoms are NMR-invisible due to extensive conformational exchange. Therefore, single point mutagenesis was used to assign three of these last four $I^{\delta1}$ or $L^{\delta2}$ resonances (I104, L107 and I110) (Fig. 3). The remaining residue, L103, was assigned by a careful re-analysis of both HCC and backbone triple resonance experiments performed after the assignment of I104.

Sequence specific assignment of A^β M^ε T^γ methyl groups of HSP90-NTD

In the previous U-[^{2}H, ^{15}N, ^{13}C], Ile-[2,3,4,4-$^2H_4$; 1,2,3,4-$^{13}C_4$; $[^{13}C^1H_3]^{\delta1}/[^{12}C^2H_3]^{\gamma2}$], Leu-[2,3,3,4-$^2H_4$; 1,2,3,4-$^{13}C_4$; $[^{13}C^1H_3]^{pro\text{-}S}/[^{12}C^2H_3]^{pro\text{-}R}$], Val-[2,3-$^2H_2$; 1,2,3-$^{13}C_3$; $[^{13}C^1H_3]^{pro-S}/[^{12}C^2H_3]^{pro-R}$] HSP90-NTD sample, only $I^{\delta1}L^{\delta2}V^{\gamma2}$ methyl groups were connected to backbone by a linear ^{13}C-chain. We did not incorporate labelled alanine in HSP90-NTD culture for the sample used to acquire the HCC experiments although it is commercially available with an optimal labelling pattern (1,2,3-$^{13}C_3$, 2-^{2}H-Ala). We recommend for future studies to incorporate labelled alanine in the culture medium to decrease the numbers of mutants required to complete the assignment. Regarding methionine and threonine residues, their assignment cannot be undertaken using HCC experiments. Indeed, the sulfur atom present on methionine residues prevents the use of HCC experimens to assign methionine methyl moieties from available backbone assignment and the Thr-[1,2,3-^{13}C, 2,3-$^2H_2$, $^{13}C^1H_3$-γ] labelled amino acid is not commercially available impeding the use of HCC 3D experiments to transfer assignment from backbone to threonine methyl groups in large proteins. Therefore, in order to assign the remaining 38 A^β M^ε and T^γ methyl groups, we decided to use a combination of both single point mutations and through space intermethyl NOE correlation peaks. To assign the remaining A^β, M^ε and T^γ methyl groups using inter-methyl NOE-connectivities a

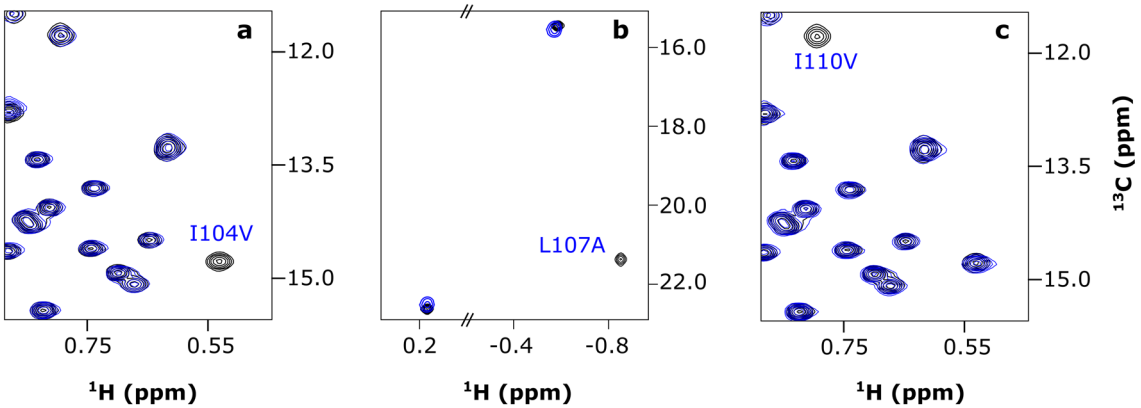


Fig. 3 Assignment of HSP90-NTD methyl groups belonging to the flexible loop covering ATP binding site. The 2D SOFAST methyl TROSY spectra were recorded using either the isoleucine to valine mutant samples using U-[^{2}H, ^{12}C, ^{15}N]-Ile-$[^{13}C^1H_3]^{\delta1}$ labelling scheme, or the leucine to alanine mutant sample using U-[^{2}H, ^{12}C, ^{15}N]-Leu/Val-$[^{13}C^1H_3]^{pro\text{-}S}$ labelling scheme. Spectra were recorded at 298 K on a NMR spectrometer operating at a proton frequency of 850 MHz. **a** The HSP90-NTD mutant spectra of I104V. **b** L107A. **c** I110V. Each mutant spectrum extract (dark blue) was superimposed with the wild type protein extract (black)

U-[^{2}H, ^{15}N, ^{12}C], Ala-[^{13}C^{1}H$_3$]$^{\beta}$, Met-[^{13}C^{1}H$_3$]$^{\varepsilon}$, Leu/Val-[^{13}C^{1}H$_3$]$^{pro-S}$, Ile-[^{13}C^{1}H$_3$]$^{\delta1}$, Thr-[^{13}C^{1}H$_3$]$^{\gamma}$ HSP90-NTD sample was required. Indeed, intermethyl NOE connectivities between previously assigned isoleucine, leucine and valine residues and unassigned A$^{\beta}$, M$^{\varepsilon}$ and T$^{\gamma}$ methyl groups simplify the assignment of methionine, threonine and alanine methyl moieties.

First small-scale samples with only one type of methyl group labelled A$^{\beta}$, M$^{\varepsilon}$ or T$^{\gamma}$ were prepared to identify the amino acids type corresponding to each correlation remaining to assign in the 2D methyl-TROSY spectrum. Then, we overexpressed and purified in small scale 30 single point ^{13}C^{1}H$_3$-labelled mutants. In order to minimize secondary chemical shifts replacement amino acids that were structurally similar to the substituted amino acid were chosen (Crublet et al. 2014) (Fig. S4).

Each of these samples containing 100 to 500 µg of ^{13}C^{1}H$_3$-labelled single point mutant of HSP90-NTD were used to acquire a 2D SOFAST-methyl-TROSY spectrum (Amero et al. 2009, 2011) on a NMR spectrometer operating at a ^{1}H frequency of 850 MHz and equipped with a sample changer and a 1.7 mm cryogenic probe head. This mutant library allowed us to assign unambiguously 27 methyl groups. Three spectra of single point mutants were more complex to analyze due to chemical shift perturbations that result from the introduced mutation (Fig. S5). Out of the 38 A$^{\beta}$, M$^{\varepsilon}$ and T$^{\gamma}$ methyl groups, eleven remained unassigned after the first analysis of the mutant library, among them eight methyls (one M$^{\varepsilon}$, two A$^{\beta}$ and five T$^{\gamma}$ methyl groups) for which mutants were not available and three T$^{\gamma}$ methyl groups whose mutant spectra were challenging to analyze. The sole unassigned methionine methyl signal was unambiguously assigned as the last remaining methionine residue (M98).

To complete the assignment for the last 10 methyl groups, a 0.5 mM sample of U-[^{2}H, ^{15}N, ^{12}C], Ala-[^{13}C^{1}H$_3$]$^{\beta}$, Met-[^{13}C^{1}H$_3$]$^{\varepsilon}$, Leu/Val-[^{13}C^{1}H$_3$]^{proS}, Ile-[^{13}C^{1}H$_3$]$^{\delta1}$, Thr-[^{13}C^{1}H$_3$]$^{\gamma}$ labelled HSP90-NTD was prepared to acquire a 3D HMQC-NOESY-HMQC. A total of 344 intermethyl NOEs cross peaks with a S/N ≥ 5 were detected between methyls distant by up to 10 Å. Examples of intermethyl NOE and a matrix presenting all the observed NOEs are displayed in Fig. 4. To assign the remaining 2 alanines and 8 threonines, the NOE connectivities and the previously assigned methyls (49 I$^{\delta1}$L$^{\delta2}$V$^{\gamma2}$ and 28 A$^{\beta}$ M$^{\varepsilon}$ and T$^{\gamma}$ methyl groups) were used as input for the program MAGIC. Six methyl groups [2 alanines (A141 and A145) and 4 threonines (T90, T88, T115 and T149)] were unambiguously assigned with both a high percentage of NOE correlation peaks assigned and a high confidence score. Four threonines (T94, T109, T176 and T184) were left either with multiple assignments or an assignment with a low confidence score. However, taking into account the information obtained from the NOE-based assignment, the mutant spectra displaying chemical shift perturbations (Fig. S5) were carefully reanalyzed allowing us to assign the four remaining threonine methyl signals.

Finally, all the 87 methyl groups of HSP90-NTD were assigned using both isoleucine precursor and the newly labelled acetolactate for isoleucine, leucine and valine methyl groups and this mixed NOE/mutants strategy for alanine, methionine and threonine methyl moieties (Fig. 5, Table S2). It can be noted that, there is, indeed, an overlap of two isoleucines (I33/I128) explaining only 86 visible peaks. Even though dynamics in the intermediate regime broaden the backbone resonances of the segment which is covering the ATP binding site [103–115] beyond the detection

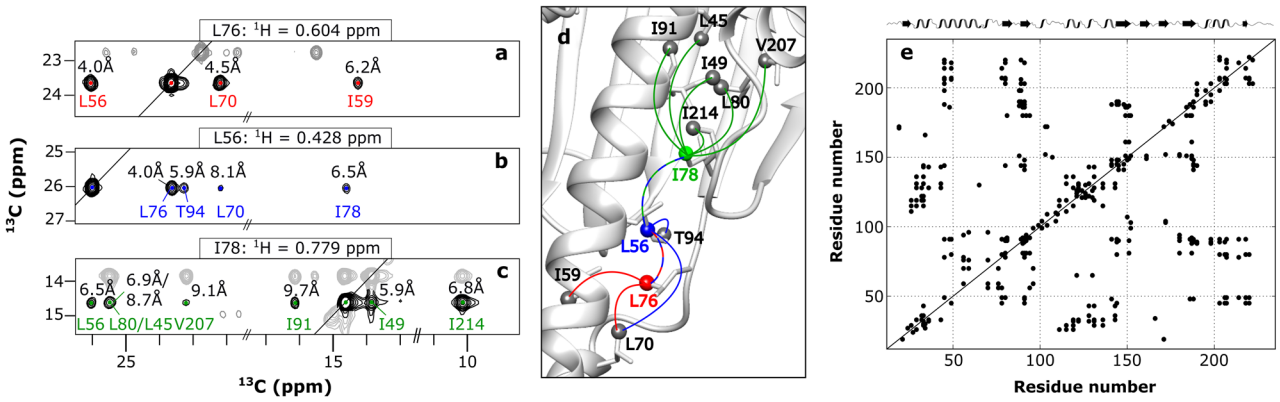


Fig. 4 Detected intermethyl NOEs in human HSP90-NTD. **a–c** Examples of 2D extracts of a 3D HMQC-NOESY-HMQC experiment recorded using U-[^{2}H, ^{12}C, ^{15}N]-Leu/Val-[^{13}C^{1}H$_3$]$^{pro-S}$, Ile-[^{13}C^{1}H$_3$]$^{\delta1}$, Met-[^{13}C^{1}H$_3$]$^{\varepsilon}$, Ala-[^{13}C^{1}H$_3$]$^{\beta}$, Thr-[^{13}C^{1}H$_3$]$^{\gamma}$ HSP90-NTD sample on a NMR spectrometer operating at a proton frequency of 950 MHz. The planes were extracted at the methyl proton frequencies of L76 (**a**), L56 (**b**) and I78 (**c**). The NOEs detected are colored in red, blue and green, respectively. **d** The NOEs detected in 2D extracts presented in panels a-c are displayed on the 3D structure of HSP90-NTD (PDB: 1YES) by lines (red for L76, blue for L56 and green for I78). **e** 2D matrix representing all the HSP90-NTD methyl residue pairs for which NOE cross-peaks have been detected

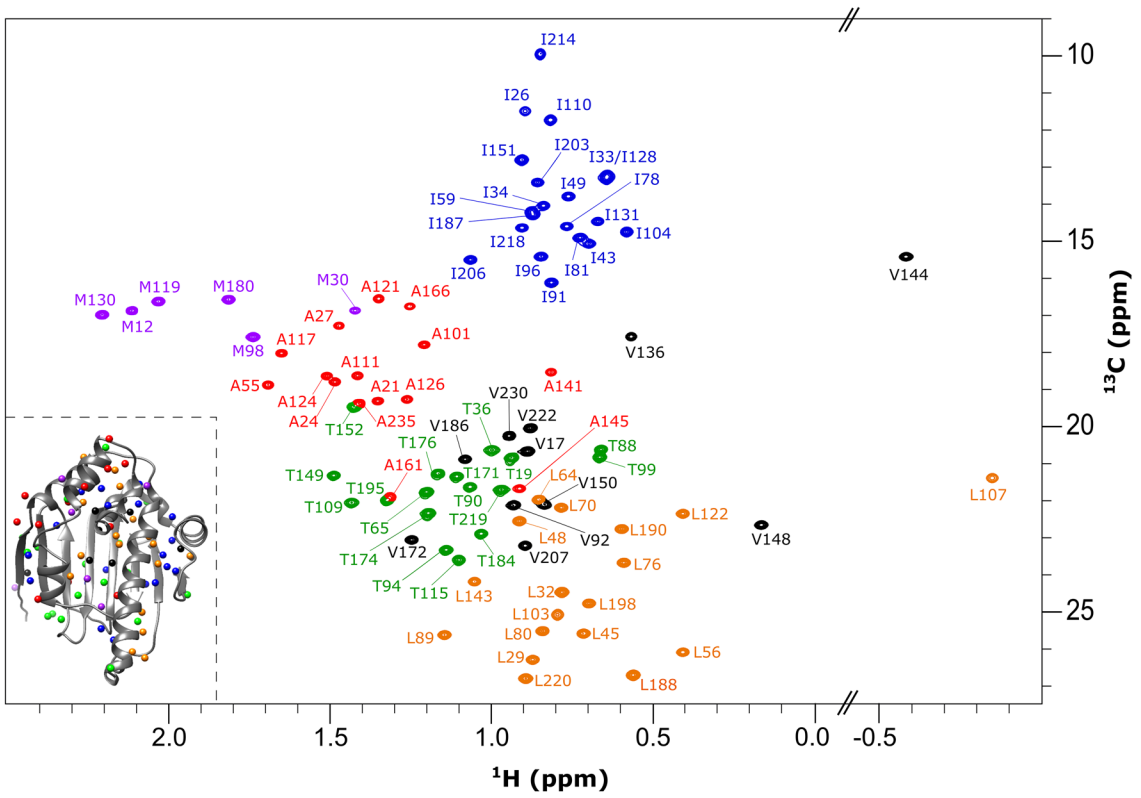


Fig. 5 Assigned 2D ^{1}H-^{13}C SOFAST methyl TROSY spectrum of apo HSP90-NTD. Human HSP90-NTD was perdeuterated and specifically $^{13}C^{1}H_3$-labelled on Leu/Val-[$^{13}C^{1}H_3$]$^{pro\text{-}S}$, Ile-[$^{13}C^{1}H_3$]$^{\delta 1}$, Met-[$^{13}C^{1}H_3$]$^{\varepsilon}$, Ala-[$^{13}C^{1}H_3$]$^{\beta}$, Thr-[$^{13}C^{1}H_3$]$^{\gamma}$ methyl groups. Each signal is annotated with the corresponding residue number. The spectrum was recorded on a NMR spectrometer operating at a proton frequency of 950 MHz. On the bottom left side an insert represents the 3D structure of human HSP90-NTD (PDB: 1YES). The methyl groups are represented by spheres. Alanines, isoleucines, valines, leucines, threonines and methionines are depicted in red, dark blue, black, orange, light green and purple, respectively

threshold, the according methyl probes are visible. The assignment of L103, I104, L107, T109, I110, A111 and T115 render this previously invisible region amenable to NMR studies.

Conclusion

This research reports the synthesis of a new dissymmetric $^{13}C/^{2}H$-labelled acetolactate, a precursor that allows to connect directly, via linear ^{13}C chains, backbone atoms to the pro-*S* methyl groups of leucine and valine residues. This optimized precursor can be combined with isoleucine precursor and $^{2}H/^{13}C$-alanine to enable the transfer of assignment from backbone to methyl groups with only one sample without requiring the correction of $^{1}H/^{2}H$ isotopic chemical shift for the ^{13}C resonances. We expect that this new precursor will ease the assignment of leucine and valine pro-*S* methyl groups of proteins using already assigned backbone resonances as it simplifies the analysis of the NMR experiments. This innovative labelling scheme was applied to the 29 kDa N-terminal domain of human HSP90 protein and to the 87 kDa hetero hexameric prefoldin complex. Using both isoleucine precursor and the newly labelled acetolactate, we managed to simply and efficiently transfer the backbone sequential assignment to all the isoleucine-δ_1, leucine and valine pro-*S* methyl moieties of HSP90-NTD. This allowed us to confirm or correct the residue specific assignment of most isoleucine, leucine and valine methyl groups (2 assignments were corrected for the 49 ILV residues—Table S2) and the stereospecific assignment of prochiral methyl groups (5 stereospecific assignments were inverted for the 29 leucine and valine residues—Table S2). In addition to the full assignment of $I^{\delta 1}L^{\delta 2}V^{\gamma 2}$ methyl groups, we have used a mixed strategy based on mutagenesis and intermethyl NOEs to assign 38 new methyl resonances corresponding to the $A^{\beta}M^{\varepsilon}T^{\gamma}$ methyl moieties of HSP90-NTD. Hence, we show that, despite extended conformational exchange that impedes the complete backbone assignment, we managed to detect and assign signals for all methyl probes including the ones belonging to the segment covering HSP90 ATP binding site.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10858-021-00370-0.

Acknowledgements The authors thank Dr. R Awad and Mr. L. Imbert for advice and stimulating discussions. This work used the high field NMR and isotopic labelling facilities at the Grenoble Instruct-ERIC Center (ISBG; UAR 3518 CNRS-CEA-UGA-EMBL) within the Grenoble Partnership for Structural Biology (PSB). Platform access was supported by FRISBI (ANR-10-INBS-05-02) and GRAL, a project of the University Grenoble Alpes graduate school (Ecoles Universitaires de Recherche) CBH-EUR-GS (ANR-17-EURE-0003). IBS acknowledges integration into the Interdisciplinary Research Institute of Grenoble (IRIG, CEA). This work was supported by grants from CEA/NMR-Bio (research program C24990) and by the French National Research Agency in the framework of the "*Investissements d'avenir*" program (ANR-15-IDEX-02).

Data availability The FIDs acquired for this study are available in the biological magnetic resonance databank (bmrbig12) and the assignment have been deposited under the BMRB ID: 50786.

References

Amero C, Schanda P, Asunción Durá M et al (2009) Fast two-dimensional NMR spectroscopy of high molecular weight protein assemblies. J Am Chem Soc 131:3448–3449. https://doi.org/10.1021/ja809880p

Amero C, Asunción Durá M, Noirclerc-Savoye M et al (2011) A systematic mutagenesis-driven strategy for site-resolved NMR studies of supramolecular assemblies. J Biomol NMR 50:229–236. https://doi.org/10.1007/s10858-011-9513-5

Ayala I, Sounier R, Usé N et al (2009) An efficient protocol for the complete incorporation of methyl-protonated alanine in perdeuterated protein. J Biomol NMR 43:111–119. https://doi.org/10.1007/s10858-008-9294-7

Ayala I, Hamelin O, Amero C et al (2012) An optimized isotopic labelling strategy of isoleucine-γ2 methyl groups for solution NMR studies of high molecular weight proteins. Chem Commun 48:1434–1436. https://doi.org/10.1039/c1cc12932e

Ayala I, Chiari L, Kerfah R et al (2020) Asymmetric synthesis of methyl specifically labelled l-threonine and application to the NMR studies of high molecular weight proteins. ChemistrySelect 5:5092–5098. https://doi.org/10.1002/slct.202000827

Chao F-A, Kim J, Xia Y et al (2014) FLAMEnGO 2.0: an enhanced fuzzy logic algorithm for structure-based assignment of methyl group resonances. J Magn Reson 245:17–23. https://doi.org/10.1016/jjmr.2014.04.012

Crublet E, Kerfah R, Mas G et al (2014) A cost-effective protocol for the parallel production of libraries of CH3-specifically labeled mutants for NMR studies of high molecular weight proteins. Methods Mol Biol 1091:229–243. https://doi.org/10.1007/978-1-62703-691-7_17

Delaglio F, Grzesiek S, Vuister GW et al (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6:277–293. https://doi.org/10.1007/BF00197809

Elif Karagöz G, Duarte AMS, Ippel H et al (2011) N-terminal domain of human Hsp90 triggers binding to the cochaperone p23. Proc Natl Acad Sci USA 108:580–585. https://doi.org/10.1073/pnas.1011867108

Epstein J, Cannon P, Swidler R, Baraze A (1977) Amplification of cyanide ion production by the micellar reaction of keto oximes with phosphono- and phosphorofluoridates. J Org Chem 42:759–762. https://doi.org/10.1021/jo00424a043

Favier A, Brutscher B (2019) NMRlib: user-friendly pulse sequence tools for Bruker NMR spectrometers. J Biomol NMR 73:199–211. https://doi.org/10.1007/s10858-019-00249-1

Gans P, Hamelin O, Sounier R et al (2010) Stereospecific isotopic labeling of methyl groups for NMR spectroscopic studies of high-molecular-weight proteins. Angew Chemie - Int Ed 49:1958–1962. https://doi.org/10.1002/anie.200905660

Gardner KH, Kay LE, Chinchilla D, Fisher K (1997) Production and incorporation of 15 N, 13 C, 2 H (1 H-δ1 methyl) isoleucine into proteins for multidimensional NMR studies. J Am Chem Soc 119:7599–7600

Gauto DF, Estrozi LF, Schwieters CD et al (2019) Integrated NMR and cryo-EM atomic-resolution structure determination of a half-megadalton enzyme complex. Nat Commun 10:2697. https://doi.org/10.1038/s41467-019-10490-9

Gelis I, Bonvin AMJJ, Keramisanou D et al (2007) Structural basis for signal sequence recognition by the 204-kDa translocase motor SecA determined by NMR. Cell 131:756–769. https://doi.org/10.1016/j.cell.2007.09.039

Goto NK, Gardner KH, Mueller GA et al (1999) A robust and cost-effective method for the production of Val, Leu, Ile (δ1) methyl-protonated 15 N-, 13 C-, 2 H-labeled proteins. J Biomol NMR 13:369–374

Gross JD, Gelev VM, Wagner G (2003) A sensitive and robust method for obtaining intermolecular NOEs between side chains in large protein complexes. J Biomol NMR 25:235–242

Hajduk PJ, Augeri DJ, Mack J et al (2000) NMR-based screening of proteins containing 13 C-labeled methyl groups. J Am Chem Soc 122:7898–7904. https://doi.org/10.1021/ja000350l

Isaacson RL, Simpson PJ, Liu M et al (2007) A new labeling method for methyl transverse relaxation-optimized spectroscopy NMR spectra of alanine residues. J Am Chem Soc 129:15428–15429. https://doi.org/10.1021/ja0761784

Jacobs DM, Langer T, Elshorst B et al (2006) NMR backbone assignment of the N-terminal domain of human HSP90. J Biomol NMR 36:52

Kerfah R, Hamelin O, Boisbouvier J, Marion D (2015a) CH3-specific NMR assignment of alanine, isoleucine, leucine and valine methyl groups in high molecular weight proteins using a single sample. J Biomol NMR 63:389–402. https://doi.org/10.1007/s10858-015-9998-4

Kerfah R, Plevin MJ, Pessey O et al (2015b) Scrambling free combinatorial labeling of alanine-β, isoleucine-δ1, leucine-proS and valine-proS methyl groups for the detection of long range NOEs. J Biomol NMR 61:73–82. https://doi.org/10.1007/s10858-014-9887-2

Kerfah R, Plevin MJ, Sounier R et al (2015c) Methyl-specific isotopic labeling: a molecular tool box for solution NMR studies of large proteins. Curr Opin Struct Biol 32:113–122. https://doi.org/10.1016/j.sbi.2015.03.009

Lapinaite A, Simon B, Skjaerven L et al (2013) The structure of the box C/D enzyme reveals regulation of RNA methylation. Nature 502:519–523. https://doi.org/10.1038/nature12581

Lescanne M, Skinner SP, Blok A et al (2017) Methyl group assignment using pseudocontact shifts with PARAssign. J Biomol NMR 69:183–195. https://doi.org/10.1007/s10858-017-0136-3

Lescanne M, Ahuja P, Blok A et al (2018) Methyl group reorientation under ligand binding probed by pseudocontact shifts. J Biomol NMR 71:275–285. https://doi.org/10.1007/s10858-018-0190-5

Mas G, Crublet E, Hamelin O et al (2013) Specific labeling and assignment strategies of valine methyl groups for NMR studies of high molecular weight proteins. J Biomol NMR 57:251–262. https://doi.org/10.1007/s10858-013-9785-z

Mas G, Guan J-Y, Crublet E et al (2018) Structural investigation of a chaperonin in action reveals how nucleotide binding regulates

the functional cycle. Sci Adv 4:eaau:4196. https://doi.org/10.1126/sciadv.aau4196
Monneau YR, Rossi P, Bhaumik A et al (2017) Automatic methyl assignment in large proteins by the MAGIC algorithm. J Biomol NMR 69:215–227. https://doi.org/10.1007/s10858-017-0149-y
Nerli S, De Paula VS, McShan AC, Sgourakis NG (2021) Backbone-independent NMR resonance assignments of methyl probes in large proteins. Nat Commun 12:691. https://doi.org/10.1038/s41467-021-20984-0
Ohtaki A, Kida H, Miyata Y et al (2008) Structure and molecular dynamics simulation of archaeal prefoldin: the molecular mechanism for binding and recognition of nonnative substrate proteins. J Mol Biol 376:1130–1141. https://doi.org/10.1016/j.jmb.2007.12.010
Park SJ, Kostic M, Dyson HJ (2011) Dynamic interaction of Hsp90 with its client protein p53. J Mol Biol 411(1):158–173. https://doi.org/10.1016/j.jmb.2011.05.030
Pritišanac I, Degiacomi MT, Alderson TR et al (2017) Automatic assignment of methyl-NMR spectra of supramolecular machines using graph theory. J Am Chem Soc 139:9523–9533. https://doi.org/10.1021/jacs.6b11358
Pritišanac I, Würz JM, Alderson TR, Güntert P (2019) Automatic structure-based NMR methyl resonance assignment in large proteins. Nat Commun 10:4922. https://doi.org/10.1038/s41467-019-12837-8
Pritišanac I, Alderson TR, Güntert P (2020) Automated assignment of methyl NMR spectra from large proteins. Prog Nucl Magn Reson Spectrosc 118–119:54–73. https://doi.org/10.1016/j.pnmrs.2020.04.001
Rosenzweig R, Moradi S, Zarrine-Afsar A et al (2013) Unraveling the mechanism of protein disaggregation through a ClpB-DnaK interaction. Science 339:1080–1083. https://doi.org/10.1126/science.1233066
Sounier R, Blanchard L, Wu Z, Boisbouvier J (2007) High-accuracy distance measurement between remote methyls in specifically protonated proteins. J Am Chem Soc 129:472–473. https://doi.org/10.1021/ja067260m
Sprangers R, Kay LE (2007) Quantitative dynamics and binding studies of the 20S proteasome by NMR. Nature 445:618–622. https://doi.org/10.1038/nature05512
Stoffregen MC, Schwer MM, Renschler FA, Wiesner S (2012) Methionine scanning as an NMR tool for detecting and analyzing biomolecular interaction surfaces. Structure 20:573–581. https://doi.org/10.1016/j.str.2012.02.012
Törner R, Awad R, Gans P et al (2020) Spectral editing of intra- and inter-chain methyl–methyl NOEs in protein complexes. J Biomol NMR 74:83–94. https://doi.org/10.1007/s10858-019-00293-x
Törner R, Henot F, Awad R et al (2021) Backbone and methyl resonances assignment of the 87 kDa prefoldin from Pyrococcus horikoshii. Biomol NMR Assignment. https://doi.org/10.1007/s12104-021-10029-4
Tugarinov V, Kay LE (2003) Ile, Leu, and Val methyl assignments of the 723-residue malate synthase G using a new labeling strategy and novel NMR methods. J Am Chem Soc 125:13868–13878. https://doi.org/10.1021/ja030345s
Tugarinov V, Hwang PM, Ollerenshaw JE, Kay LE (2003) Cross-correlated relaxation enhanced 1H-13 C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. J Am Chem Soc 125:10420–10428. https://doi.org/10.1021/ja030153x
Tugarinov V, Kay LE (2004a) Stereospecific NMR assignments of prochiral methyls, rotameric states and dynamics of valine residues in malate synthase G. J Am Chem Soc 126:9827–9836. https://doi.org/10.1021/ja048738u
Tugarinov V, Kay LE (2004b) An isotope labeling strategy for methyl TROSY spectroscopy. J Biomol NMR 28:165–172
Tugarinov V, Kay LE, Ibraghimov I, Orekhov VY (2005) High-resolution four-dimensional 1H-13 C NOE spectroscopy using methyl-TROSY, sparse data acquisition, and multidimensional decomposition. J Am Chem Soc 127:2767–2775. https://doi.org/10.1021/ja044032o
Velyvis A, Ruschak AM, Kay LE (2012) An economical method for production of 2H,13CH3-threonine for solution NMR studies of large protein complexes: application to the 670 kDa proteasome. PLoS ONE 7:e43725. https://doi.org/10.1371/journal.pone.0043725
Vranken WF, Boucher W, Stevens TJ et al (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. Proteins Struct Funct Genet 59:687–696. https://doi.org/10.1002/prot.20449
Xu Y, Matthews S (2013) MAP-XSII: an improved program for the automatic assignment of methyl resonances in large proteins. J Biomol NMR 55:179–187. https://doi.org/10.1007/s10858-012-9700-z